Mutational analysis of 4-coumarate:CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes

Hans-Peter Stuible, Daniela Büttner[1], Jürgen Ehlting[2], Klaus Hahlbrock, Erich Kombrink*

Max-Planck-Institut für Züchtungsforschung, Abteilung Biochemie, Carl-von-Linné-Weg 10, 50829 Köln, Germany

Received 30 November 1999; received in revised form 8 January 2000

Edited by Ulf-Ingo Flügge

Abstract 4-Coumarate:coenzyme A ligase (4CL) is a key enzyme of general phenylpropanoid metabolism which provides the precursors for a large variety of important plant secondary products, such as lignin, flavonoids, or phytoalexins. To identify amino acids important for 4CL activity, eight mutations were introduced into *Arabidopsis thaliana* At4CL2. Determination of specific activities and K_m values for ATP and caffeate of the heterologously expressed and purified proteins identified four distinct classes of mutants: enzymes with little or no catalytic activity; enzymes with greatly reduced activity but wild-type K_m values; enzymes with drastically altered K_m values; and enzymes with almost wild-type properties. The latter class includes replacement of a cysteine residue which is strictly conserved in 4CLs and had previously been assumed to be directly involved in catalysis. These results substantiate the close relationship between 4CL and other adenylate-forming enzymes such as luciferases, peptide synthetases, and fatty acyl-CoA synthetases.

Key words: Phenylpropanoid metabolism; Luciferase; Peptide synthetase; Fatty acyl-CoA synthetase; *Arabidopsis thaliana*

1. Introduction

4-Coumarate:CoA ligase (4CL; EC 6.2.1.12) catalyzes the activation of 4-coumarate and various other cinnamic acid derivatives to the corresponding thiol esters in a two-step reaction via an adenylate intermediate [1,2]. This reaction represents the last step in a short series of biochemical conversions known as general phenylpropanoid metabolism, leading from phenylalanine to the activated cinnamic acid derivatives which are precursors for the biosynthesis of a large variety of plant secondary metabolites with functions as diverse as their structures [3–5]. Phenylpropanoids serve as structural components (lignin, suberin and other cell wall-associated phenolics), protectants against biotic and abiotic stresses (phytoalexins, antioxidants, UV-absorbing compounds), flower pigments (flavonoids, anthocyanins), and signal molecules (salicylic acid, nodulation factors). Because of its functional importance, the general phenylpropanoid pathway has been extensively studied in a large variety of plants, in particular with respect to regulation of the corresponding genes and properties of the encoded proteins [3–7]. However, this does not include detailed studies concerning the reaction mechanism catalyzed by 4CL. Earlier experimental approaches to identify functionally relevant amino acids were severely hampered by the instability of the mutant proteins expressed in *Escherichia coli* [2]. Thus, all available information about catalytically important amino acid positions is based on sequence comparison or indirect evidence obtained from chemical modifications of the protein [2,8,9].

The presence of several conserved peptide motifs in all available 4CL amino acid sequences has been repeatedly observed by computer assisted sequence alignments [2,8,10–12]. The Box I motif, SSGTTGLPKGV, is not only almost absolutely conserved in 4CLs, but highly similar motifs are also found in luciferases, acetyl-CoA synthetases, long-chain fatty acyl-CoA synthetases and peptide synthetases [13,14]. The presence of this putative nucleotide-binding motif has even been used as one important criterion to establish a superfamily of adenylate-forming enzymes [13]. In the presence of ATP and Mg^{2+} all members of this enzyme superfamily form an adenylate intermediate which is esterified with either CoA (4CL, acetyl-CoA synthetases, acyl-CoA synthetases) or the enzyme-bound CoA derivative 4′-phosphopantheteine (peptide synthetases), or oxidized by molecular oxygen (luciferases). The Box II motif, GEICIRG, is absolutely conserved in all 4CLs, and its central cysteine residue has been suggested to be directly involved in catalysis [2]. Inhibition experiments with sulfhydryl-modifying agents which abolished 4CL activity supported such a function [9], which was also in good agreement with the originally proposed model for the reaction mechanism of multienzymes such as non-ribosomal peptide synthetases [15,16]. However, evidence in conflict with the participation of a cysteine residue in catalysis accumulated for other adenylate-forming enzymes, resulting in a revised model of adenylate formation by peptide synthetases, which excluded the direct involvement of a cysteine residue in this reaction step [17,18].

We recently cloned and functionally expressed in *E. coli* three genes encoding 4CL isoforms of *Arabidopsis thaliana* [12]. The high activity and stability of the expressed proteins prompted us to investigate the functional importance of selected amino acids previously proposed to be essential for

*Corresponding author. Fax: (49)-221-5062313.
E-mail: kombrink@mpiz-koeln.mpg.de

[1] Present address: Martin-Luther Universität, Institut für Genetik, Weinbergweg 22, 06108 Halle, Germany.
[2] Present address: Department of Botany, University of British Columbia, Vancouver, B.C. V6T 1Z4, Canada.

Abbreviations: 4CL, 4-coumarate:coenzyme A ligase; ATP, adenosine 5′-triphosphate; CoA, coenzyme A; IPTG, isopropyl-β-D-thiogalactoside; PCR, polymerase chain reaction

PII: S0014-5793(00)01133-9

enzymatic activity. A series of mutations was introduced into the At4CL2 reading frame, and the purified, mutated proteins were analyzed for their kinetic properties with respect to the substrates caffeate and ATP. Based on these data, the reaction mechanism previously postulated for 4CLs is disproved and the close relationship between 4CL and other adenylate-forming enzymes, such as peptide synthetases, luciferases and fatty acyl-CoA synthetases, is experimentally substantiated.

2. Materials and methods

2.1. Bacterial strains, plasmids, primer sequences and DNA manipulation techniques

Standard DNA manipulation techniques were performed as previously described [19]. For plasmid amplification *E. coli* strains XL1-Blue (Stratagene, La Jolla, USA) or DH5a (GibcoBRL Life Technologies, Rockville, USA) were used. Protein expression was performed using *E. coli* strain M15[pRP4] (Qiagen, Hilden, Germany). The pQE30-based At4CL2 expression plasmid used for mutagenesis experiments has previously been described [12]. Site-directed in vitro mutagenesis of the At4CL2 GEICIRG motif was achieved by replacement of the respective DNA region by synthetic oligonucleotide adapters. For this purpose the pBluescript derivative containing the At4CL2 cDNA as a *Bam*HI/*Kpn*I fragment was digested with *Avr*II and *Msc*I, resulting in deletion of the GEICIRG coding region. The deleted DNA fragment was replaced with a compatible oligonucleotide adapter carrying the desired mutation. After amplification, the modified At4CL2 constructs were inserted into the pQE30 expression vector using the *Bam*HI and *Kpn*I restriction sites. One strand each of the adapters used in this study is listed in Table 1. All other mutations were introduced into At4CL2 by PCR amplification of the entire double-stranded At4CL2 expression plasmid using two mutated oligonucleotide primers, each complementary to opposite strands of the vector. All components necessary for this mutagenesis procedure were included in the commercial QuikChange™ kit (Stratagene, La Jolla, USA). One primer sequence of each pair of complementary mutagenesis primers is listed in Table 1. To facilitate the identification of the mutated plasmids, a new restriction site was introduced together with the point mutation whenever possible (Table 1). The entire DNA sequences of all mutated reading frames were determined on ABI Prism 377 DNA sequencers (PE Applied Biosystems, Foster City, USA). Sequence alignments were generated using the BestFit, Gap, or PileUp programs of the GCG software package, version 10.0 [20].

2.2. Expression and purification of At4CL2 proteins

All At4CL2 coding constructs were expressed in *E. coli* strain M15[pRP4] using the expression vector pQE30 (Qiagen, Hilden, Germany) and thus contained a His_6 tag at their N-termini. For At4CL2 purification, *E. coli* transformants were grown overnight in 50 ml LB medium with 60 mg/l kanamycin and 100 mg/l ampicillin at 37°C. Cells were harvested by centrifugation and suspended in 400 ml of fresh medium without kanamycin. To induce expression, cells were grown for 7 h at 37°C in the presence of 5 mM isopropyl-1-thio-β-D-galactopyranoside, collected by centrifugation and stored at −80°C. For protein purification, pelleted cells were suspended in 80 ml of 50 mM sodium phosphate, pH 7.5, containing 300 mM NaCl (buffer A), lysed by sonification with a Branson sonifier, and insoluble components removed by centrifugation (20 000 × *g*, 20 min, 4°C). One ml of Ni-NTA agarose (Qiagen, Hilden, Germany) equilibrated with buffer A was added to the supernatant, and for adsorption of the His_6-tagged proteins the suspension was stirred on ice for 1 h. The Ni-NTA agarose was washed twice with 40 ml of buffer A and three times with 40 ml of 50 mM sodium phosphate, pH 7.5, 300 mM NaCl, 30 mM imidazole, 15% glycerol (buffer B). Subsequently, the agarose was packed to a column and the protein eluted by stepwise increase from 0.1 to 0.5 M imidazole dissolved in buffer B. Fractions (1 ml) were collected and aliquots analyzed by SDS–polyacrylamide gel electrophoresis which was performed as previously described [21]. After addition of dithiothreitol to a final concentration of 2 mM, the At4CL2-containing fractions were stored at −80°C.

2.3. Enzyme assays

4CL activity was determined with the spectrophotometric assay as previously described, using caffeic acid as the preferred phenolic substrate [9,12]. The change in absorbance was monitored at 363 nm, the absorption maximum for caffeoyl-CoA [22]. The K_m values for caffeate were estimated at fixed concentrations of ATP (5.5 mM) and CoA (0.3 mM) by linear regression of *v*/*s* against *s* (Hanes plot), and the K_m values for ATP at fixed concentrations of caffeate (0.2 mM) and CoA (0.3 mM). Since V_{max} values could not be determined for all mutant enzymes, specific activity was determined under standard conditions (0.2 mM caffeate, 5.5 mM ATP, 0.3 mM CoA). Protein concentrations were determined according to Bradford with bovine serum albumin as standard [23].

3. Results and discussion

3.1. Site-directed mutagenesis and heterologous expression of At4CL2

For in vitro mutagenesis a recently described expression plasmid was used in which the At4CL2 reading frame was fused to a N-terminal His_6 tag and placed under the control of a strong inducible *E. coli* promoter [12]. To verify that the N-terminal tag has no influence on activity and substrate specificity of 4CL, the K_m values for ATP and caffeate – which in comparison to 4-coumarate is the preferred substrate of At4CL2 [12] – were determined for both enzyme variants, At4CL2 containing or lacking the six additional histidine residues. Both enzymes showed nearly identical K_m values of about 160 μM for ATP and 25 μM for caffeate, which is in accordance with our previous results [12] and those reported for many native, purified 4CLs [1,10,24]. Therefore, the His_6-tagged version seemed to be well suited for mutagenesis since it allowed easy purification and further protein chemical studies.

Positions for in vitro mutagenesis were chosen either for their absolute conservation within all known 4CLs or for their

Table 1
Sequences of DNA adapters and PCR primers used for the construction of At4CL2 mutant proteins

Sequence[a]	Mutation	Restriction site[b]
DNA adapters		
CTAGGAACAAACCCAACCAAATCATGAAAGGCTATCTCAATGACCCCTTGG	ΔGEICIRG	Δ*Nsi*I
CTAGGAACAAACCCGGCCAAATATGCATCCGTGGCAACCAAATCATGAAAGGCTATCTCAATGACCCCTTGG	Glu^{401}/Gln	*Eae*I
CTAGGAACAAACCCGGCGAAATAGCAATCCGTGGCAACCAAATCATGAAAGGCTATCTCAATGACCCCTTGG	Cys^{403}/Ala	Δ*Nsi*I
PCR primers		
CGACGGGTCTCCCCTCCGGAGTGATGCTAAC	Lys^{211}/Ser	*Acc*III
CGAGCTTTTCATTGTGGATCAATTGAAAGAACTCATC	Arg^{449}/Gln	*Mun*I
GATTGAAAGAACTCATCACGTACAAAGGATTTCAAG	Lys^{455}/Thr	–
GAAAGAACTCATCAAGTACTCCGGATTTCAAGTGGCTCCAGCTG	Lys^{457}/Ser	*Acc*III
CCTAAAGCTCCATCAGGGAATATTTTGAGGAAGGATCTAAGAG	Lys^{540}/Asn	*Ssp*I

[a]Modified codons are underlined.
[b]Restriction sites were introduced or deleted (Δ) to facilitate the identification of plasmids carrying the desired mutation.

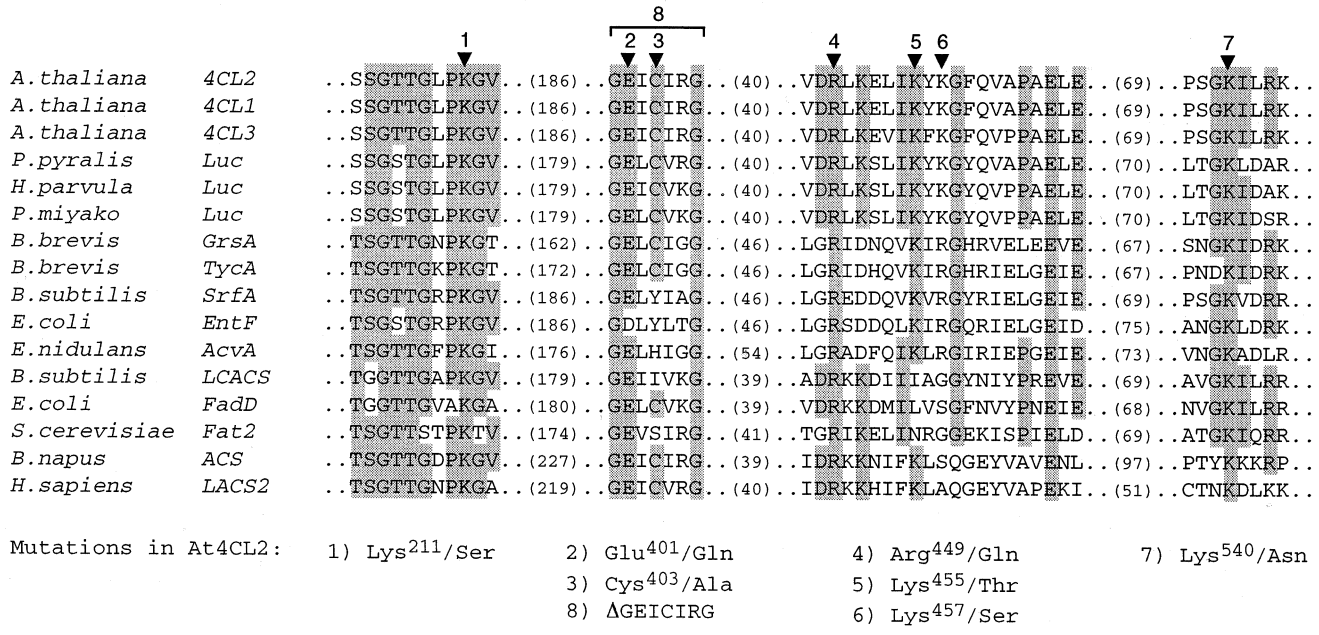

Fig. 1. Sequence comparison of conserved peptide motifs from selected members of the adenylate-forming enzyme superfamily. Amino acid residues common to at least 10 of the aligned sequences are shaded (i.e. identity > 60%). The mutations introduced in At4CL2 are marked by arrowheads, labeled with numbers, and specified below the sequences. The proteins included in the alignment are as follows (with accession numbers given in brackets): *Arabidopsis thaliana* 4CL2 (AAD47193), 4CL1 (AAD47191), 4CL3 (AAD47194); *Photinus pyralis* luciferase (P08659); *Hotaria parvula* luciferase (AAC37253); *Pyrocoelia miyako* luciferase (AAC37254); *Brevibacillus brevis* gramicidin S synthetase 1 (GrsA) (P14687); *Brevibacillus brevis* tyrocidine synthetase 1 (TycA) (P09095); *Bacillus subtilis* surfactin synthetase 1 (SrfA) (P27206); *Escherichia coli* enterobactin synthetase component F (EntF) (P11454); *Emericella nidulans* ACV synthetase (AcvA) (CAA38631); *Bacillus subtilis* long-chain acyl-CoA synthetase (LCACS) (P94547); *Escherichia coli* long-chain acyl-CoA synthetase (FadD) (A45062); *Saccharomyces cerevisiae* fatty acyl-CoA synthetase (Fat2) (P38137); *Brassica napus* acyl-CoA synthetase (ACS) (CAA64327); *Homo sapiens* long-chain fatty acyl-CoA synthetase 2 (LACS2) (AAD17853).

high conservation within the superfamily of adenylate-forming enzymes consisting of 4CLs, luciferases, peptide synthetases and long-chain fatty acyl-CoA synthetases (Fig. 1). The first set of mutations included the substitution of a cysteine residue (Cys^{403}) postulated to be directly involved in thiol ester formation by 4CLs [2]. Since multiple ion pair interactions with lysine side chains have been observed in several nucleotide-binding proteins [25], a second set of mutations included individual replacements of four lysine residues in At4CL2 by neutral amino acids (Fig. 1). After site-directed mutagenesis, the complete modified reading frames were verified by sequencing before the respective plasmids were used for protein expression. Western blot analysis of crude protein extracts from IPTG-treated *E. coli* cells showed that all mutant enzymes were expressed at approximately the same level without detectable degradation (not shown). Purification by Ni^{2+}-chelate affinity chromatography resulted in almost homogeneous protein preparations (Fig. 2) that could be stored at −80°C for several months with a low decrease in enzymatic activity ranging from 15 to 20%.

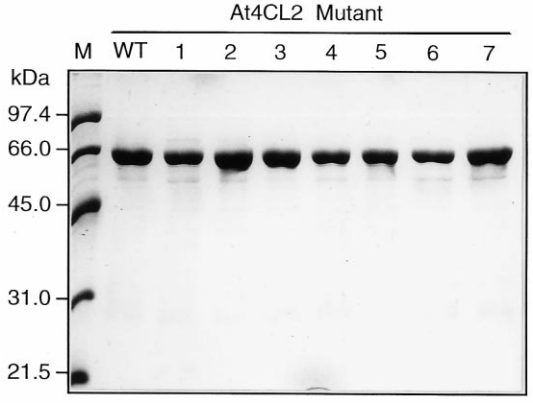

Fig. 2. SDS–polyacrylamide gel of heterologously expressed and purified At4CL2 mutant enzymes. Proteins (4 μg per lane) were separated by SDS–polyacrylamide gel electrophoresis and stained with Coomassie brilliant blue. WT, wild-type At4CL2; 1, Lys^{211}/Ser; 2, Glu^{401}/Gln; 3, Cys^{403}/Ala; 4, Arg^{449}/Gln; 5, Lys^{455}/Thr; 6, Lys^{457}/Ser; 7, Lys^{540}/Asn. M, molecular weight standard, low range (Bio-Rad Laboratories, München, Germany).

3.2. Properties of At4CL2 mutant enzymes with an altered GEICIRG motif

Two mutations introduced into the GEICIRG motif of At4CL2 (Glu^{401}/Gln, Cys^{403}/Ala) reduced the specific activity of the mutant enzymes to 21 and 45% of the wild-type level, respectively, whereas the K_m values for ATP or caffeate were not significantly altered (Table 2). Thus, since the point mutant Cys^{403}/Ala still exhibited considerable enzymatic activity, the direct involvement of Cys^{403} in adenylate or thiol ester formation can be excluded, although the reduction in activity may indicate a supportive role of this amino acid in catalysis or stabilization of the protein structure. These data are in agreement with the proposed functional relationship between 4CLs, luciferases and peptide synthetases, since substitutions of the corresponding cysteine residues in firefly luciferase and tyrocidine synthetase from *Bacillus brevis* also affected enzymatic activity only moderately [26,27]. Further support for the

conclusion that the cysteine residue of the GEICIRG motif is not essential for enzymatic activity may be derived from the fact that this amino acid is not strictly conserved in luciferases, peptide synthetases and acyl-CoA synthetases (Fig. 1).

The glutamic acid residue at position 401 in At4CL2 was chosen for substitution based on sequence comparison of numerous members of the adenylate-forming enzyme superfamily which revealed that this is the most highly conserved charged amino acid residue of the GEICIRG motif (Fig. 1). However, the substitution Glu^{401}/Gln also did not greatly affect the enzymatic properties (Table 2). A direct participation of the GEICIRG motif in the catalytic mechanism is therefore unlikely, although its strict conservation in all 4CLs may indicate an important function. Such importance is underscored by the total loss of activity in the mutant enzyme, ΔGEICIRG, from which the whole GEICIRG motif had been deleted (Table 2). However, the most likely explanation for its inactivity is a gross perturbation of the structural integrity of the enzyme due to the deletion of an important peptide motif rather than the loss of amino acids that are essential for catalysis.

3.3. At4CL2 mutant enzymes with an altered putative nucleotide-binding domain

A second highly conserved peptide motif found in all members of the superfamily of adenylate-forming enzymes [13] is the putative AMP-binding domain signature as defined in the PROSITE database (PDOC00427, [28]). This motif is rich in Gly, Ser and Thr and contains near the C-terminus an absolutely conserved Lys (Fig. 1). These structural features resemble the Walker type A motif or P loop, GXXXXGK(S/T), which is found in numerous ATP- or GTP-binding proteins [29,30] where it forms a so-called phosphate-binding loop. In this flexible loop the Lys is proposed to interact with the γ-phosphate of the nucleotide triphosphate [29,30]. Therefore, substitution of the corresponding Lys^{211} in the SSGTTGLPK^{211}G motif of At4CL2 was expected to have profound effects on enzyme activity and on the K_m for ATP. Indeed, the mutant enzyme with the Lys^{211}/Ser substitution exhibited only about 3% of wild-type activity. However, the K_m values for both, ATP and caffeate, were not significantly affected (Table 2).

Similar results were recently described for the heterologously expressed adenylation domain of tyrocidine synthetase 1 (TycA) from *Bacillus brevis*, where substitution of the conserved Lys^{198} in the TSGTTGKPK^{198}G signature motif also resulted in a strong decrease in enzyme activity, whereas K_m values and ATP-binding affinity were only slightly affected [14,27]. In contrast, substitution of either one of two lysine residues in the putative nucleotide-binding motif of surfactin (SrfA) from *Bacillus subtilis* (corresponding to Leu^{209} and Lys^{211} in At4CL2) had only moderate effects on the adenylation activity of the mutated enzyme [31]. Accordingly, it appears questionable whether the SSGTTGLPKG signature motif of At4CL2 is directly involved in nucleotide binding. A similar conclusion was recently drawn from the crystal structure of the phenylalanine-activating subunit of gramicidin synthetase 1 (PheA) in a complex with AMP and phenylalanine [32]. In this conformation, this conserved signature motif forms a flexible loop between two β-strands and was suggested to accommodate the pyrophosphate leaving group during catalysis rather than directly participating in ATP binding [32].

3.4. At4CL2 mutant enzymes with alterations in the C-terminal domain

The arginine corresponding to position 449 in At4CL2 is invariant amongst all adenylate-forming enzymes and was therefore mutated in several of these proteins. In At4CL2 the substitution Arg^{449}/Gln reduced the enzymatic activity to 3% of the wild-type level with a concomitant 10-fold increase in the K_m values for ATP and caffeate (Table 2). Substitutions of the corresponding Arg^{416} in tyrocidine synthetase 1 (TycA) from *Bacillus brevis* or Arg^{453} in fatty acyl-CoA synthetase (FadD) from *E. coli* resulted in a profound loss of catalytic efficiency in both cases, indicating that the charged arginine residue in this position is important for activity [14,33]. In tyrocidine synthetase 1, as in At4CL2, this substitution not only lowered the catalytic capacity of the mutant enzyme but also reduced its binding affinity for the acyl substrate. From the crystal structure of gramicidin synthetase 1 (PheA) it has been deduced that the invariant Arg^{428} is located in a flexible loop connecting the two subdomains of this adenylation subunit, and it was proposed to function in cooperation with the likewise conserved Lys of the SSGTTGLPKG motif (see above) in coordinating the pyrophosphate release during adenylate formation [32]. In spite of their similar putative functions, the differences observed for the Lys^{211}/Ser and Arg^{449}/Gln mutations in At4CL2 may indicate that Arg^{449} has a stronger influence on the structural integrity of the enzyme than Lys^{211}.

Two point mutations (Lys^{455}/Thr, Lys^{457}/Ser) were intro-

Table 2
Enzymatic properties of heterologously expressed At4CL2 variants

Mutant	K_m caffeate (μM)	K_m ATP (μM)	Specific activity (nkat/mg)	Specific activity (% of wild-type)
Wild-type	24	163	209	100
Lys^{211}/Ser	14	151	5.6	2.7
Glu^{401}/Gln	28	240	44	21
Cys^{403}/Ala	18	173	94	45
Arg^{449}/Gln	> 400[a]	1507	6.2	2.9
Lys^{455}/Thr	–	–	0.33	0.16
Lys^{457}/Ser	5	29	7.7	3.7
Lys^{540}/Asn	–	–	< 0.07	–
DGEICIRG	–	–	n.d.	–

All values are the average of three independent determinations.
n.d., not detectable.
–, not determined, enzyme activity too low.
[a]Enzyme activity increased linearly up to the highest applicable caffeate concentration (400 μM).

duced into the region next to the putative flexible loop which in peptide synthetases has been proposed to be related to aminoacyl adenylation [25,34] and was recently shown to be stabilized against proteolytic cleavage upon adenylate formation [35]. The substitution Lys^{457}/Ser in At4CL2 resulted in a 96% reduction of enzymatic activity, comparable to that observed with Arg^{449}/Gln (see above), however, with the opposite effect on the K_m values for ATP and caffeate, i.e. 5-fold decreases compared to wild-type (Table 2). Lys^{457} is not strictly conserved amongst adenylate-forming enzymes; it is present in 4CLs and luciferases, whereas in peptide synthetases the corresponding position is occupied by an arginine residue and in fatty acyl-CoA synthetases no positively charged residue is found. Therefore, it seems unlikely that Lys^{457} is directly involved in adenylate formation. However, the unusual properties of the Lys^{457}/Ser mutant, i.e. decreased K_m values, indicate that at least in 4CL this position plays an important role in coordinating the sequential reaction steps or in product release. Substitution of the neighboring lysine residue (Lys^{455}/Thr) resulted in reduction of the catalytic activity to less than 0.2% (Table 2). This very low activity prevented the determination of reliable kinetic constants, but estimates seem to indicate decreases of the K_m values for ATP and caffeate, as observed above for the Lys^{457}/Ser substitution. Protective affinity labeling of tyrocidine synthetase 1 by fluorescein 5′-isothiocyanate (FITC) identified the conserved Lys^{422}, corresponding to Lys^{455} in At4CL2, as one of two residues associated with ATP binding [25]. Our enzymatic data support the notion that the equivalent of Lys^{455} is functionally important in most adenylate-forming enzymes, with the possible exception of bacterial and yeast fatty acyl-CoA synthetases in which this position is not conserved.

The second labeled lysine residue of tyrocidine synthetase 1 corresponds to Lys^{540} in At4CL2. The Lys^{540}/Asn substitution has the most drastic effect of all point mutations introduced into the At4CL2 reading frame, with enzymatic activity being almost completely abolished, again preventing the determination of kinetic constants (Table 2). Substitution of the corresponding Lys^{945} in surfactin (SrfA) resulted in a similar, drastically reduced activity [31]. These results are in accordance with conclusions drawn from the crystal structure of the adenylation subunits of gramicidin synthetase 1, indicating close electrostatic interactions of Lys^{517}, corresponding to Lys^{540} of At4CL2, with both substrates, the phosphate moiety of AMP and the carboxyl group of the activated amino acid phenylalanine [32]. It has therefore been suggested that this lysine residue stabilizes the negatively charged penta-valent transition state [32] and, being absolutely conserved in this enzyme superfamily, functions as the active site of adenylate formation.

4. Conclusions and outlook

The results presented in this study provide the first experimental evidence on the functional importance of conserved amino acid residues and motifs of 4CL and thereby not only confirm a close relationship between 4CL and other adenylate-forming enzymes, but also disprove a direct involvement of the conserved GEICIRG motif of 4CLs in their catalytic activity. In addition, our data lay the foundation for further detailed biochemical studies. The high-level expression of mutated 4CL2 variants, their stability as active enzymes even during long-term storage, the simple optical enzyme assay, the available sequence information of a large number of 4CLs from different plants, render 4CL2 from *Arabidopsis* a suitable system to study the steric requirements or constraints determining differential utilization/conversion of various cinnamic acid derivatives as described for numerous 4CL isoforms [12]. The identification of individual amino acids affecting the substrate specificity of 4CL is not only of academic interest. The position of 4CL at a metabolic branch point connecting general phenylpropanoid metabolism with the different endproduct-specific pathways generating a myriad of biologically important compounds, makes 4CL a particularly suitable target for pathway or product engineering. Designing enzymes with altered properties could give rise to new or modified combinations of phenylpropanoid compounds in plants [36,37]. To unravel the molecular basis of substrate specificity has been attempted for various adenylate-forming enzymes, in particular acyl-CoA synthetases and peptide synthetases [33,38], and at least in some cases, a correlation between the polarity of the acyl substrate and some amino acid residues forming the hypothetical substrate pocket could be revealed [32]. However, the ultimate goal of designing novel substrate specificities has not yet been achieved and progress is possibly impeded by the large sizes and complex, modular compositions of these enzymes. The relatively low structural complexity and ease of handling suggest that 4CL may serve as a suitable model for such studies.

Acknowledgements: We thank Drs. Imre E. Somssich and Wolfgang Knogge for critical reading of the manuscript and Roswitha Lentz for excellent technical assistance. H.P.S. is recipient of a post-doctoral fellowship of the Deutsche Forschungsgemeinschaft (Stu 251/1-1).

References

[1] Knobloch, K.-H. and Hahlbrock, K. (1975) Eur. J. Biochem. 52, 311–320.
[2] Becker-André, M., Schulze-Lefert, P. and Hahlbrock, K. (1991) J. Biol. Chem. 266, 8551–8559.
[3] Hahlbrock, K. and Grisebach, H. (1979) Annu. Rev. Plant Physiol. 30, 105–130.
[4] Hahlbrock, K. and Scheel, D. (1989) Annu. Rev. Plant Physiol. Plant Mol. Biol. 40, 347–369.
[5] Dixon, R.A. and Paiva, N.L. (1995) Plant Cell 7, 1085–1097.
[6] Douglas, C.J. (1996) Trends Plant Sci. 1, 171–178.
[7] Weisshaar, B. and Jenkins, G.I. (1998) Curr. Opin. Plant Biol. 1, 251–257.
[8] Uhlmann, A. and Ebel, J. (1993) Plant Physiol. 102, 1147–1156.
[9] Knobloch, K.-H. and Hahlbrock, K. (1977) Arch. Biochem. Biophys. 184, 237–248.
[10] Allina, S.M., Pri-Hadash, A., Theilmann, D.A., Ellis, B.E. and Douglas, C.J. (1998) Plant Physiol. 116, 743–754.
[11] Hu, W.-J., Kawaoka, A., Tsai, C.-J., Lung, J., Osakabe, K., Ebinuma, H. and Chiang, V.L. (1998) Proc. Natl. Acad. Sci. USA 95, 5407–5412.
[12] Ehlting, J., Büttner, D., Wang, Q., Douglas, C.J., Somssich, I.E. and Kombrink, E. (1999) Plant J. 19, 9–20.
[13] Fulda, M., Heinz, E. and Wolter, F.P. (1994) Mol. Gen. Genet. 242, 241–249.
[14] Dieckmann, R., Pavela-Vrancic, M., Pfeifer, E., von Döhren, H. and Kleinkauf, H. (1997) Eur. J. Biochem. 247, 1074–1082.
[15] Lipmann, F. (1980) Adv. Microbiol. Physiol. 21, 227–260.
[16] Kleinkauf, H. and von Döhren, H. (1990) Eur. J. Biochem. 192, 1–15.
[17] Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R. and Morris, H.R. (1996) J. Biol. Chem. 271, 15428–15435.

[18] Kleinkauf, H. and von Döhren, H. (1996) Eur. J. Biochem. 236, 335–351.
[19] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[20] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[21] Kombrink, E., Schröder, M. and Hahlbrock, K. (1988) Proc. Natl. Acad. Sci. USA 85, 782–786.
[22] Stöckigt, J. and Zenk, M.H. (1975) Z. Naturforsch. 30c, 352–358.
[23] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[24] Voo, K.S., Whetten, R.W., O'Malley, D.M. and Sederoff, R.R. (1995) Plant Physiol. 108, 85–97.
[25] Pavela-Vrancic, M., Pfeifer, E., Schröder, W., von Döhren, H. and Kleinkauf, H. (1994) J. Biol. Chem. 269, 14962–14966.
[26] Ohmiya, Y. and Tsuji, F.I. (1997) FEBS Lett. 404, 115–117.
[27] Gocht, M. and Marahiel, M.A. (1994) J. Bacteriol. 176, 2654–2662.
[28] Hofmann, K., Bucher, P., Falquet, L. and Bairoch, A. (1999) Nucleic Acids Res. 27, 215–219.
[29] Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) EMBO J. 1, 945–951.
[30] Saraste, M., Sibbald, P.R. and Wittinghofer, A. (1990) Trends Biochem. Sci. 15, 430–434.
[31] Hamoen, L.W., Eshuis, H., Jongbloed, J., Venema, G. and van Sinderen, D. (1995) Mol. Microbiol. 15, 55–63.
[32] Conti, E., Stachelhaus, T., Marahiel, M.A. and Brick, P. (1997) EMBO J. 16, 4174–4183.
[33] Black, P.N., Zhang, Q., Weimar, J.D. and DiRusso, C.C. (1997) J. Biol. Chem. 272, 4896–4903.
[34] Tokita, K., Hori, K., Kurotsu, T.M.K. and Saito, Y. (1993) J. Biochem. 114, 522–527.
[35] Dieckmann, R., Pavela-Vrancic, M., von Döhren, H. and Kleinkauf, H. (1999) J. Mol. Biol. 288, 129–140.
[36] Lamb, C.J., Ryals, J.A., Ward, E.R. and Dixon, R.A. (1992) Bio/Technology 10, 1436–1445.
[37] Dixon, R.A. and Steele, C.L. (1999) Trends Plant Sci. 4, 394–400.
[38] Marahiel, M.A. (1997) Chem. Biol. 4, 561–567.